A bilingual digital library for academic and entrepreneurial knowledge management

Ranka Stanković *

Faculty of Mining and Geology, University of Belgrade
Đušina 7, 11000 Belgrade, Serbia

Cvetana Krstev

Faculty of Philology, University of Belgrade
Studentski trg 3, 11000 Belgrade, Serbia

Biljana Lazić

Faculty of Mining and Geology, University of Belgrade
Đušina 7, 11000 Belgrade, Serbia

Dalibor Vorkapić

Faculty of Mining and Geology, University of Belgrade
Đušina 7, 11000 Belgrade, Serbia

** Corresponding author*

Structured Abstract

Purpose – A generic knowledge management process of organization, storage and retrieval of knowledge can suitably be fitted in a digital library. In the digital and knowledge age digital libraries can be used in knowledge management to handle intellectual assets and support knowledge creation. A multilingual digital library either stores content in more than one language or provides multilingual query access to monolingual content. In Serbia 18 of 308 scientific journals regularly published are bi-lingual, with papers simultaneously being in English and Serbian. Bilingual publication of results of international projects is also an increasingly frequent practice. In this paper we present a publicly available multilingual digital library named *Bibliša*, developed for management, search and the browsing of aligned bilingual text collections.

Design/methodology/approach – The approach to the development of the presented digital library was to store its content in a NoSQL-database, with a web tool to enable the use of rich information in the stored text collections. The library content originates from various bilingual sources, with all documents aligned at the sentence level and provided with metadata. In designing *Bibliša* special attention was given to its component that offers language support, based on bilingual lexical resources and tools, to various multilingual aspects of this library.

Originality/value – Users have a high level of flexibility in searching the library content: they can search the library using metadata, or perform full-text search by keywords of

their choice. Furthermore, the user's original query, issued in Serbian or English, can be expanded to the other language, both morphologically and semantically. Thus it offers a novel access to digital content to its users. In addition to that, *Bibliša* presents an original approach to successful combining of several components: Lexical resources, Library content, Web services and Web interface.

Practical implications– Digital library users often need to transform information access into knowledge creation. To that end, *Bibliša* offers numerous documents from several scientific journals and projects covering various domains. Metadata and full-text search are available within a user-friendly platform, with the possibility for a very high level of search refinement. *Bibliša* also supports the analysis of concepts and concept structures, identifying terms assigned to concepts, and establishing relations between terms in English and Serbian, which makes it a valuable tool for terminological research.

Keywords – Knowledge management, Digital library, Multilingualism, Language Resources, Terminology

Paper type – Academic Research Paper

1 Introduction

According to Lesk, "a digital library is a collection of information that is both digitized and organized" (Lesk 2005; p. 2); therefore, "a digital library can be searched for any phrase; it can be accessed all over the world; and it can be copied without error". Abram sees the process for knowledge creation and use as a continuum where data transforms into information, information transforms into knowledge and knowledge drives and underpins behaviour and decision-making (Abram,1997; p.20). To deal with the almost infinite amount of digital information, new methods such as data mining, text mining, content management, search engines, spidering programs, natural language searching, linguistic analysis, semantic networks, knowledge extraction, etc. should be a part of recent developments in knowledge management systems.

DL researchers have a need to transform information access into knowledge creation. Instead of serving as information providers, DL could become knowledge repositories by effectively categorizing, analyzing and organizing the contents of digital libraries. Lexical and terminological resources allow DL to provide more useful services. (Chen, 2004).

Having in mind that a terminology research consists of: analyzing the concepts and concept structures used in a field or domain of activity, identifying the terms assigned to the concepts and in the case of bilingual or multilingual terminology, establishing correspondences between terms in the various languages, we have made an effort to develop a software tool and bi-lingual resource that supports terminology research. This paper presents a developed solution, named *Bibliša*[1], which is free for use and publicly available.

[1] *http://hlt.rgf.bg.ac.rs/Biblisha/*

In the second section of this paper are presented the aspects of a multilingual digital libraries and motivation for our research and software implementation. The Third section is dedicated to design and methodology. It starts with discussion about relations between knowledge management and digital library, followed by introduction of *Bibliša* and a few example of usage. Furthermore, in the fourth section we explain system components: lexical resources, text collection as library content and web services and application. Searching capabilities are given in the fifth section, comprising description of full text search by keywords and metadata search. Finally, in conclusion we summarise main features and purpose of the developed system, as well as further ideas about improvement.

2 Multilingualism in DL

There are two aspects of a multilingual DL – it can be a library "with content in more than one language", or a library that "provides multilingual query access to monolingual content" (Diekema, 2012; p. 166). Multilingual libraries can also differ in respect to the number of languages included, the type and domain of documents they collect, the geographical region they cover, audiences they address, and the financial and institutional support they get.

Digital libraries with multilingual content can be further differentiated. Most of them offer various works in different languages with no intention of providing the same works in several languages. It does not come as a surprise that there are not many digital libraries that offer the same texts in two or more languages as their development is not straightforward: various versions of texts have to be collected and adequately prepared.

In recent years new types of multilingual (or rather bi-lingual) journals have emerged in which academic and professional papers are published in two languages. In Serbia, a small European country of approximately 7 million inhabitants, 18 journals out of 308 scientific journals regularly published are bi-lingual. Journals are published bilingually for two main reasons: the results of Serbian scientists are presented to a much wider audience if they are published in English and not only in Serbian and the journal can attract scientists from wider geographic regions, which can improve its rating.

More and more international projects in Serbia publish research results bilingually in Serbian and English, so that the initial concept of a library of bilingual journals is extended to a more general collection of bilingual documents. All TEMPUS[1] projects have an obligation to publish reports and results bilingually, including the BAEKTEL project[2] (Blending Academic and Entrepreneurial Knowledge in Technology enhanced learning). In this paper we will illustrate the capabilities of the developed system by aligned bi-lingual deliverables of the BAEKTEL project.

[1] *http://erasmusplus.rs/category/education/*
[2] *http://www.baektel.eu/*

3 Design and methodology

3.1. Knowledge management and digital library

Roknuzzaman reported overlaps between digital libraries and knowledge management, stating that a generic knowledge management process of acquisition, organization, storage and retrieval, and dissemination of knowledge with receiving feedback can suitably be fitted in a digital library (Roknuzzaman et al. 2009). The integration of knowledge management can add value to developing a knowledge-based culture, management of intellectual assets, promotion of knowledge sharing, innovations in digital library services. Knowledge assets and their relation to performance improvement of an enterprise thus became an important research area in the management of company business (Linzalone, 2008).

Jain and Mutula stated that library and information professionals have been referred to as knowledge managers; and libraries and information centres, have been referred to as centres of knowledge (Jain and Mutula, 2008).

Wen reports that of all types of libraries, special libraries, especially business and corporate libraries, are taking the lead on Knowledge Management research; and of academic libraries, public services and digital libraries are in the limelight (Wen, 2005). Within academic libraries, public services are taking the lead in the research and implementation of Knowledge Management.

It is a well known fact that knowledge is recognized as the ultimate resource of our days while digital libraries provide extraordinary information and knowledge highways for all. Lytras outlines the need for building library systems that go a step beyond the existing technological infrastructure and present an attempt to clarify the main directions for research on the concept of a "semantic digital library", and the main management and technical challenges derived from such an idea/ideas (Lytras, 2005).

The overall requirements for a semantic approach to digital libraries are use of ontologies, lexical and terminological resources.

3.2. Introducing Bibliša

We have developed a DL and a web tool, dubbed *Bibliša,* so that both direct and indirect users, accessing the abundant information from journals published bilingually, can get the most use out of (Stanković et al., 2012a). Its content are documents from various journals and project reports that are published in two languages – Serbian and English. All documents are provided with the usual metadata (article's author(s), publication date, title, etc.) and are aligned at the sentence level.

Besides searching by metadata, *Bibliša* offers a full-text search by keywords of the user's choice. A user's original query, which can be issued in each of the two languages, can be expanded in several different ways: by translating into other languages, morphologically by including all inflected forms of keywords in a query, and semantically by adding synonyms or other related words. The query expansion is supported by various language resources, described in section 4.1.

Additionally, *Bibliša* represents a valuable tool for terminological research. For languages with unstable terminology it is an indispensable tool for authors not only when searching but also when writing their own articles as it can help them make appropriate terminological choices.

Finally, the modular architecture of *Bibliša* enables its modules to be used in other tools (such as web service access in lexical resources, access to aligned resources, etc.)

4 System components

In designing *Bibliša* special attention is given to its language support component. It supports various aspects of multilingual libraries: its content is not only multilingual, but also aligned and it can be searched in any language. The proposed tool basically consists of the following components:

- *Lexical Resources*–used to enhance and refine users' queries;
- *Library content*–documents in two languages aligned at the sentence level;
- *Web services* – used to access lexical resources;
- *Web interface* – used for users' access to the library content.

Figure 1 presents components of our DL system and the interaction between them. When the user requests search by keyword, *Bibliša* calls web service *Vebran* to find synonyms and other related terms if requested, as well as their translations. After the user refines the offered list of terms (e.g. by deleting unneeded ones), *Bibliša* calls web services for morphological query expansion, forms an http request in the XQuery language form and sends it to the MarkLogic database. The result is presented in the XHTML form.

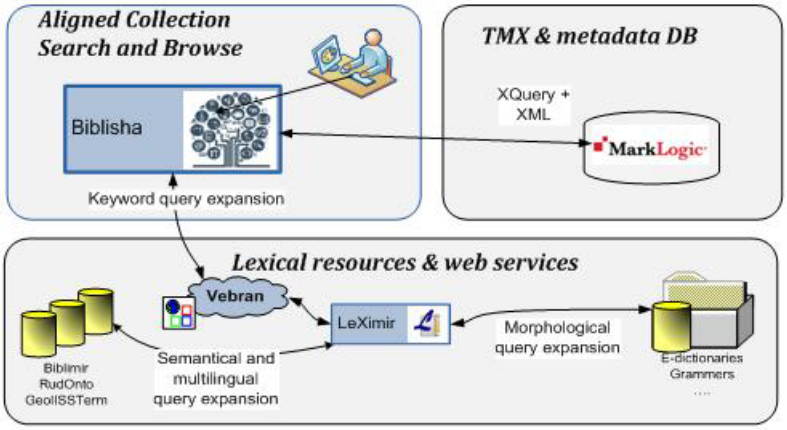


Figure 1: Bibliša: the system's components

4.1 Lexical resources

Lexical Resources are used to enhance and refine users' queries. The query expansion is supported by e-dictionaries (Serbian morphological e-dictionaries), general purpose semantic networks (English and Serbian WordNets) and domain terminological bases and ontologies. A Dictionary of Library and Information Sciences[1] is a terminology dictionary that covers the library and information sciences and related disciplines in Serbian, English and German developed at the National Library of Serbia. GeolISSTerm and RudOnto are developed at the Faculty of Mining and Geology, University of Belgrade (Stanković et al., 2012b). GeolISSTerm[2] is a thesaurus of geological terms with entries in Serbian and English, developed within the GeolISS[3] project. RudOnto[4] is another complex terminological resource, also being developed at the Faculty of Mining and Geology with the aim of becoming the future e-format reference resource in Serbian for mining terminology. Currently RudOnto comprises terms in Serbian, their English equivalents, and a small number of equivalents in other languages.

Specific features of Serbian grammar, especially its rich morphology, need corresponding language resources in the form of morphological e-dictionaries and grammars, implemented by finite state automata, finite state transducers and multi-word unit inflection rules (Krstev, 2008). All of them are used to improve the full-text search performance of our DL, especially the recall.

4.2 Library content

The DL presented in this paper presently offers numerous articles from scientific journals covering domains of library and information sciences, digital humanities, architecture and mining and also project reports on academic and entrepreneurial knowledge and e-learning.

It contains articles from three journals INFOtheca - Journal for Digital Humanities, Underground Mining Engineering, and Architecture and Urbanism. INFOtheca collection contains 74 papers from 11 issues published from 2007-2014, Underground Mining Engineering has 55 papers in 8 issues, dated from 2004-2013, while Architecture and Urbanism presently has just 10 papers. Project reports included originate from the BAEKTEL Tempus project.

All bilingual texts are aligned at the sentence level, represented in a TMX[5] (Translation Memory eXchange) format, and stored in the MarkLogic NoSQL database. Text collections and metadata are stored in MarkLogic that is a novel database management system for large scale non-structured and semi-structured data. MarkLogic is

[1] *http://rbi.nb.rs/en/dict.html*
[2] *http://geoliss.mre.gov.rs/term/*
[3] *Geological information system of Serbia (GeolISS) was financed by the Ministry of Environmental Protection during the period 2004–2006, and is presently maintained by Ministry of Mining and Energy. Many institutions working with geological data are using GeolISS.*
[4] *http://rudonto.rgf.bg.ac.rs/ is the system of ontologies developed for the mining engineering and its applications in mining equipment and mine safety domains.*
[5] *http://www.gala-global.org/oscarStandards/tmx/tmx14b.html*

a document database that has evolved from a native XML DBMS database to enterprise NoSQL. In one platform, it combines a database, search engine and application services.

The preliminary alignment phase consists of preparing an XML document (eXtensible Markup Language) according to TEI (Text Encoding Initiative) guidelines.[1] Practically, this means that the main divisions of a text as well as its titles, paragraphs and segments (sentences) have to be XML tagged. Any text editing software with support for well-formedness checking and validation according to a DTD (Document Type Definition) or XML Scheme can be used for that purpose. The next key step is the alignment itself: the task is to establish relations between translation equivalents in both texts. In this case, segments are paired that usually represent whole sentences or some of their parts. Software XAlign and Concordancier, developed in Loria labaratory in France (Laboratoire Lorrain de Recherche en Informatique et ses Applications) are used for alignment. The alignment method is based on the number of characters (length of the segment). Utvić reports that this approach is very successful (as much as 96% correctly paired documents). XAlign is now integrated into Unitex[2], a corpus processing system, based on automata-oriented technology (Utvić et al., 2007).

Text preparation, alignment and generation of TMX documents are done within a special-purpose tool ACIDE (Aligned Corpora Integrated Development Environment) (Utvić et al., 2007). The TMX document consists of TU[3] (Translation Unit) and TUV (Translation Unit Variant) elements, where each TUV is a segment in one of the languages. The following example illustrates a single aligned segment (TU) of a document:

```
<tu><proptype="Domain">BAEKTEL 1.1, 2014, ID: 5.2014.1.1</prop>
  <tuvxml:lang="en" creationid="n1 " creationdate="20150327T125605Z">
    <seg>Analysis of Existing European Practice and Principles in Development of
    Open Educational Resources (OER) </seg>
  </tuv>
  <tuvxml:lang="sr" creationid="n1 " creationdate="20150327T125605Z">
    <seg>Analiza postojeće evropske prakse i principa u razvoju otvorenih obrazovnih
    resursa (OER) </seg>
  </tuv>
</tu>
```

4.3 Web services and application

Web services are used to access lexical resources and perform query expansion: multilingual, semantical and morphological. A Web interface is offered to the user for the metadata and full-text search by entering queries in one of two languages.

---

[1] *TEI – Text Encoding Initiative: http://www.tei-c.org/index.xml*

[2] *Unitex – Corpus Processing System: http://igm.univ-mlv.fr/~unitex/. The concept of this software was born at LADL (Laboratoired'Automatique Documentaire et Linguistique), under the direction of its director, Maurice Gross.*

[3] *http://xml.coverpages.org/tmxSpec971212.html*

For a full-text search the user formulates the initial query as one or more keywords (simple or multiword). By default, the system will search all collections, but the user can restrict it by selecting a particular collection from the dropdown list of available collections. Resources that will be used for query expansion are automatically selected by the system depending on the domain of the selected text collection (their box will be checked).

The System retrieves terms that match the given keywords from the lexical resources of a query language and then finds their equivalents in another language based on inter-lingual relations established in the lexical resources. After refinement of a query (e.g. deleting or adding terms manually), the system performs semantic and multilingual query expansion.

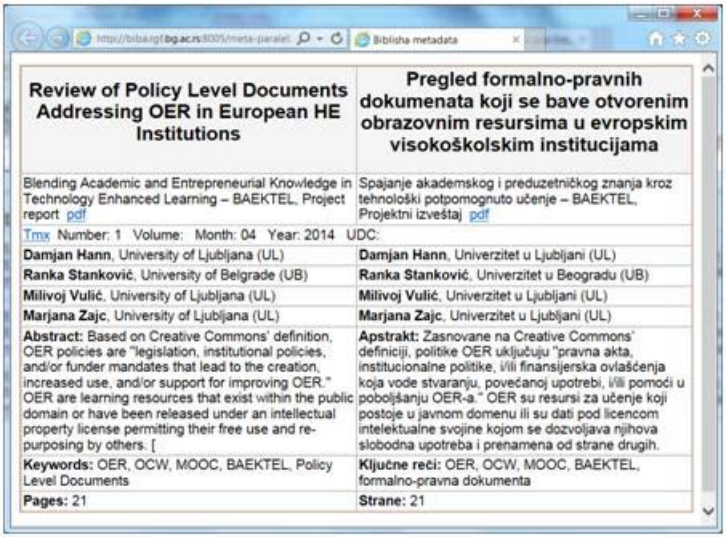
If the user so specifies, *Bibliša* forwards this query for further morphological expansion, based on morphological e-dictionaries, the system of rules for multi-word inflection, and finite automata and transducers (Krstev et al, 2008; Krstev 2008). An example of query expansion will be presented in section 5.1.

Metadata is assigned to each article in our DL. This metadata is also in the form of an XML document with a schema that consists of a set of elements designed for description of each text collection. Metadata are stored in Serbian and in English, where the attribute *lang* specifies the language of metadata element. Metadata elements provide information on:

- text collections: text collection title (journal or project), ISSN[1] and URL of the collection;
- issue for journal or work package for project: number, volume, month and year of publication;
- article for journal or deliverable for project: authors names, affiliations, e-mails, article title, categorization, pagination, abstract, keywords, URL and UDC[2].

Metadata of all collections can be browsed and searched. Browsing the metadata in Serbian and English for the selected project (Baektel) is presented in Figure 2. The user can also access the article in a TMX and PDF format following the links on this page. Further, he/she can open the page with basic data on the selected document in Serbian and English.

[1] *International Standard Serial Number*
[2] *Universal Decimal Classification number*



Review of Policy Level Documents Addressing OER in European HE Institutions	Pregled formalno-pravnih dokumenata koji se bave otvorenim obrazovnim resursima u evropskim visokoškolskim institucijama
Blending Academic and Entrepreneurial Knowledge in Technology Enhanced Learning – BAEKTEL, Project report pdf	Spajanje akademskog i preduzetničkog znanja kroz tehnološki potpomognuto učenje – BAEKTEL, Projektni izveštaj pdf
Tmx Number: 1 Volume: Month: 04 Year: 2014	UDC:
Damjan Hann, University of Ljubljana (UL)	**Damjan Hann**, Univerzitet u Ljubljani (UL)
Ranka Stanković, University of Belgrade (UB)	**Ranka Stanković**, Univerzitet u Beogradu (UB)
Milivoj Vulić, University of Ljubljana (UL)	**Milivoj Vulić**, Univerzitet u Ljubljani (UL)
Marjana Zajc, University of Ljubljana (UL)	**Marjana Zajc**, Univerzitet u Ljubljani (UL)
Abstract: Based on Creative Commons' definition, OER policies are "legislation, institutional policies, and/or funder mandates that lead to the creation, increased use, and/or support for improving OER." OER are learning resources that exist within the public domain or have been released under an intellectual property license permitting their free use and re-purposing by others. [	**Apstrakt:** Zasnovane na Creative Commons' definiciji, politike OER uključuju "pravna akta, institucionalne politike, i/ili finansijerska ovlašćenja koja vode stvaranju, povećanoj upotrebi, i/ili pomoći u poboljšanju OER-a." OER su resursi za učenje koji postoje u javnom domenu ili su dati pod licencom intelektualne svojine kojom se dozvoljava njihova slobodna upotreba i prenamena od strane drugih.
Keywords: OER, OCW, MOOC, BAEKTEL, Policy Level Documents	**Ključne reči:** OER, OCW, MOOC, BAEKTEL, formalno-pravna dokumenta
Pages: 21	**Strane:** 21

Figure 2: Bibliša: a result of the metadata search

5 Searching capabilities

5.1 Full text search by keywords

The user enters the keyword and optionally selects a text collection to search (the default is all collections). Besides the keyword itself, it is necessary to choose the keyword language, and then click on the "Preview and modify terms for query" link. The system uses web services to find synonyms and translations in selected lexical resources. The user can remove terms from the offered list or add them (a comma "," has to be used as a separator).

The system connects all the synonyms and translations stated in the panel for one language by the disjunction operator (OR). For example, for the keyword "learning", WordNet offers the following synonyms in English: "acquisition", "encyclopaedism", "encyclopaedism", "eruditeness", "erudition", "learnedness", "learning", "scholarship", while RudOnto offers just "learning". If the user wants to use other relations as well, for example hyponyms, from WordNet the system will add to this list: "carry-over", "committal to memory", "conditioning", "developmental learning", "digestion", "education", "imprinting", "incorporation", "internalisation", "internalization", "language learning", "letters", "memorisation", "memorization", "scholarship", "study", "transfer", "transfer of training", "work" and from the RudOnto: "analysis", "application", "comprehension", "evaluation", "knowledge", "synthesis". Language equivalents (translations) are listed for another language as well; in our example, for the list for Serbian will include: "sticanje znanja", "učenje", "analiza", "evaluacija", "primena", "razumevanje", "sinteza", "učenje", "znanje".

The final form of the query is obtained by the morphological expansion of each individual term, if requested by the user. For example, the original term “sticanje znanja” is expanded with inflectional forms “sticanja znanja”, “sticanju znanja”, “sticanjem znanja”, “sticanjima znanja”.

The system expands the query semantically, morphologically and in other languages, and then transforms it into XQuery[1], which is then used to search collection(s) of TMX documents. The XQuery form of request is forwarded to the TMX document database supported by the MarkLogic system. One small part of the XQuery form for the above query is presented in the following example:

http://biba.rgf.bg.ac.rs:8005/wsQ4.xqy? qJCID=5&request=

<**request**>
<query xml:lang='sr'>znanja</query>OR<queryxml:lang='sr'>znanjem</query>OR
<queryxml:lang='sr'>znanjima</query>...
<query
xml:lang='en'>acquisition</query>OR<queryxml:lang='en'>education</query><queryx
ml:lang='en'>synthesis</query>...
<operator>OR</operator>
</**request**>

As a result, aligned concordances with all the keywords in both languages are produced. At the beginning of each concordance line is the identification of the document from which the line originates – the identification is the hyperlink to a metadata card for both languages. The Metadata card has hyperlinks to TMX and PDF documents in both languages. Figure 3 shows three resulting lines for the above example query in the BAEKTEL collection.

BAEKTEL 1.3, 2014, ID: 5.2014.1.3 metadata	**n65 :** In order to promote a ***knowledge***-based society the following areas should be enhanced: establishing the links between theory and practice, including the dissemination and use of ICT, investing in research and development; strengthening the concept of lifelong ***learning***; improving flexibility and strengthening the ***education*** system, including better links between universities and the business world.	**n65 :** Područja koja treba unaprediti da se promoviše društvo zasnovano na ***znanju***: ustanoviti veze između teorije i prakse, uključujući širenje i upotrebu ICT, investiranje u istraživanje i razvoj; ojačati koncept celoživotnog ***učenja***; poboljšati fleksibilnost i jačanje obrazovnog sistema, uključujući bolje veze između univerziteta i poslovnog sveta.
BAEKTEL 1.3, 2014, ID: 5.2014.1.3 metadata	**n273 :** The commitment that Western Balkan countries channel their economic development towards ***knowledge*** and innovation in the EU accession process requires the ***application*** of new technologies to strengthen the institutional mechanisms for the development of OER.	**n273 :** Opredeljenje da zemlje Zapadnog Balkana usmeravaju ekonomski razvoj ka ***znanju*** i inovacijama u procesu pristupanju EU nalaže ***primenu*** novih tehnologija i jačanje institucionalnih mehanizama za razvoj OER-a.
BAEKTEL 1.3, 2014, ID: 5.2014.1.3 metadata	**n277 :** The development and advancement of ***knowledge*** requires greater investment in professional educational training and fostering of the implementation of innovative technologies.	**n277 :** Razvoj i unapređenje ***znanja*** zahteva veća ulaganja u obrazovno stručno usavršavanje i podsticanje ***primene*** inovativnih tehnologija.

Figure 3: An example of multi-lingual concordances

[1] *XQuery is a language for finding and extracting elements and attributes from XML documents.*

5.2 Metadata search

It is also possible to search and browse a collection by metadata in order to analyse the whole content of a collection. The search of the Serbian or English part of the collection using an author's name, words in the title, year of publishing and keywords (either individually or by combining some or all of them) is available by clicking the "Metadata search" link in the menu. To search the metadata, the user can select the language and the collection (it is possible to simultaneously search through all available text collections). The user enters the search criteria in the search field, then adds additional criteria by clicking the "+" sign. Boolean operators "OR" and "AND" build the search query. "AND" is used for different search field categories e.g. author's names and year of publication, while "OR" is used for the same search category e.g. two author's names field. The example in Figure 4 shows how the system retrieves all articles from all collections published in 2012 or 2014, whose authors are "Ranka" or "Cvetana" and the result is presented in English.

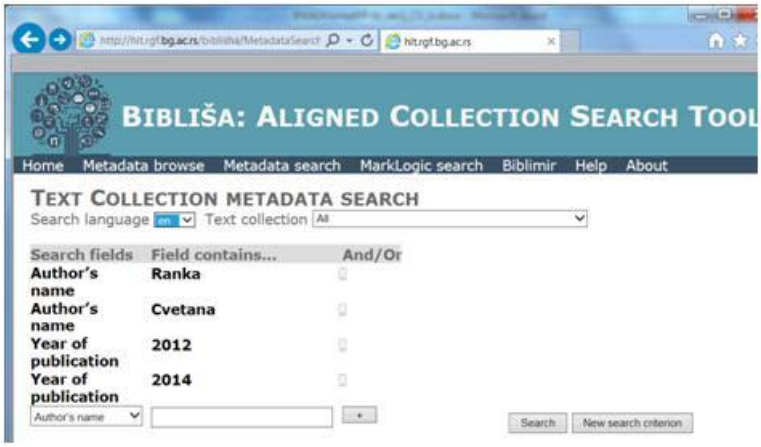


Figure 4: Bibliša: metadata search

The result of the metadata search is presented as a list of documents matching the metadata query with links to full-text articles in a PDF format, as well as to a TMX aligned text in an HTML format. Figure 5 shows the result of the previous query for the selected language (English), but selection of language "sr" would retrieve a page with metadata in Serbian.

Document	Authors	Title	Keywords
1.2012.1.1 tmx pdf	Maciej Ogrodniczuk Radovan Garabik Svetla Koeva Cvetana Krstev Piotr Pezik Tibor Pinter Adam Przepiorkowski Gyorgy Szaszak Marko Tadić Tamas Varadi Duško Vitas	Central and South-European language resources in Meta-Share	
2.2012.20.4 tmx pdf	Aleksandra Tomašević Ljiljana Kolonja Ivan Obradović Ranka Stanković Olivera Kitanović	Using UML CASE Tools for Development of an Open Pit ArcGIS Geodatabase	GIS, geodatabase, UML, ArcGIS, open pit
5.2014.1.1 tmx pdf	Damjan Hann Ranka Stanković Nenad Stefanović Snežana Šćepanović Saša Tatar Milivoj Vulić Marjana Zajc	Analysis of Existing European Practice and Principles in Development of Open Educational Resources (OER)	OER, OCW, MOOC, BAEKTEL
5.2014.1.2 tmx pdf	Damjan Hann Ranka Stanković Milivoj Vulić Marjana Zajc	Review of Policy Level Documents Addressing OER in European HE Institutions	OER, OCW, MOOC, BAEKTEL, Policy Level Documents

Figure 5: Results of the metadata search

5.3 An example of the use of Bibliša for terminology research

We will illustrate how *Bibliša* can be successfully used to establish right term equivalences. We start our search of all DL collections with one English term, "e-learning". Out of five lexical resources integrated in *Bibliša,* it is only in RudOnto that is adequate entry found, but without any Serbian equivalence. We enhance the search with a direct translation in Serbian "e-učenje". As a result, we obtain 84 concordance lines from 8 different documents: 5 articles for the journal INFOtheca and 3 Baektel deliverables. The inspection of these lines reveals that the Serbian term "e-učenje" always has as its equivalence the English term "e-learning", while the English term e-learning is translated in most cases with "elektronsko učenje" (43), but also with "e-učenje" (38), "e-obrazovanje" (2), "e-nastava" (2), and "učenje na daljinu" (6). Translation for "e-learning" occurs twice while "e-learning" appears 93 times in 84 concordances.

In the next iteration we enhance our query with these new Serbian terms. Now we obtain 112 concordance lines from the same set of documents. All new lines come from the term "učenje na daljinu" that has "distance learning" as its English equivalent in most of the cases, but also "distance education".

In the third step we add "distance learning" and "distance education" to our query. The number of concordance lines is now 123 that come from 11 documents: 8 articles for the journal INFOtheca and 3 Baektel deliverables. We discover that "distance learning" has as its equivalent, besides "učenje na daljinu" which is used most frequently, also "daljinsko učenje", "obrazovanje na daljinu", "daljinsko obrazovanje" and "udaljeno učenje".

In the fourth iteration we add these newly discovered terms to our query. The number of concordance lines remains the same – 123, which means that these Serbian terms have no other English equivalent in our DL other than distance learning. Finally, we also add electronic learning to our query, but we obtain the same result.

As a result of our research, we can conclude that the English term "e-learning" has two Serbian equivalents: "e-učenje" and "elektronsko učenje". Two more terms are

occasionally used “e-nastava” and “e-obrazovanje” but their equivalence to “e-learning” has still to be confirmed. The English term “distance learning” has Serbian equivalents “učenje na daljinu” (by far the most frequent), “daljinsko učenje”, “obrazovanje na daljinu”, “daljinsko obrazovanje” and “udaljeno učenje”. The experiment shows that the Serbian “učenje na daljinu” cannot be used as an equivalent for English e-learning.

6 Conclusions

In this paper we presented *Bibliša*, a system that offers a novel access to multilingual content to its users. Not only can they perform both a metadata and full-text search by using one user-friendly platform, but they can sophistically refine their search. The search performance improvement relies on several terminological bilingual resources and morphological e-dictionaries. Additionally, *Bibliša* represents a valuable tool for terminological research. For languages with unstable terminology, it is an indispensable tool for authors not only when searching but also when writing their own articles as it can help them make appropriate terminological choices. In order to enhance its usability in the future we plan to add to *Bibliša’s* contents bilingual manuals, tutorials, books, examples of good practice, etc. The integration and development of lexical resources during searches is being envisaged.

References

Abram, S. (1997) “Post Information Age Positioning for Special Librarians: Is Knowledge Management the Answer?”, Information Outlook, June 1997, pp. 20-21.

Diekema, A. R. (2012) "Multilinguality in the digital library: a review", Electronic Library, The 30.2: 165-181.

Hsinchun Chen, (2004) "Digital library research in the US: an overview with a knowledge management perspective", Program, Vol. 38 Iss: 3, pp.157 – 167.

Jain, P. and Mutula S. (2008) “Libraries as a learning organisations: implications for knowledge management”, Library Hi Tech News, No 8, pp. 10-14.

Krstev, C., (2008). Processing of Serbian – Automata, Texts and Electronic dictionaries. Faculty of Philology, University of Belgrade, Belgrade.

Krstev, C., Stanković, R., Vitas, D., Obradović, I. (2008). The Usage of Various Lexical Resources and Tools to Improve the Performance of Web Search Engines, in Proceedings of the 6th International Conference on Language Resources and Evaluation, LREC 2008, N. Calzolari et al., Marrakech, Morocco.

Lesk, M. (2005) Understanding digital libraries. Elsevier, San Francisco.

Linzalone, R. (2008), “Leveraging knowledge assets to improve new product development performances”, Measuring Business Excellence 12(2), pp. 38-50.

Lytras, M., Sicilia, M. A., Davies, J., & Kashyap, V. (2005), “Digital libraries in the knowledge era: Knowledge management and Semantic Web technologies”, Library Management, 26(4/5), pp. 170-175.

Roknuzzaman, Md. ,Hideaki Kanai and Katsuhiro Umemoto, (2009) "Integration of knowledge management process into digital library system: A theoretical perspective", Library Review, Vol. 58 Iss: 5, pp. 372 – 386.

Stanković, R., Krstev, C., Obradović, I., Trtovac, A., Utvić, M. (2012a), A Tool for Enhanced Search of Multilingual Digital Libraries of E-journals, in Proceedings of the 8th International Conference on Language Resources and Evaluation, LREC 2012, May 2012, Istanbul, Turkey, eds. N. Calzolari et al., European Language Resources Association Istanbul, Turkey, 2012, pp. 1710-1717, ISBN 978-2-9517408-7-7, http://www.lrec-conf.org/proceedings/lrec2012/pdf/375_Paper.pdf.

Stanković, R., Obradović, I., Kitanović, O., & Kolonja Lj. (2012b) Building Terminological Resources in an e-Learning Environment, in Proceedings of the Third International Conference on e-Learning, eLearning-2012, ed. D. Milošević, Belgrade Metropoloitan University, Belgrade.

Utvić, M. Stanković, R.Obradović, I. (2007) Integrisano okruženje za pripremu paralelizovanog korpusa, in Zbornik radova međunarodnog simpozijuma Razlike između bosanskog/bošnjačkog, hrvatskog i srpskog jezika, ed. B. Tošović, Graz, Austria.

Wen, S. (2005) Implementing Knowledge Management in Academic Libraries: A Pragmatic Approach. Chinese Librarianship, (19).